

10-1-1992

# Trends in Statewide Reading Assessment: A Closer Look

William A. Henk

*Penn State*

Robert J. Rickelman

*University of North Carolina*

Follow this and additional works at: [https://scholarworks.wmich.edu/reading\\_horizons](https://scholarworks.wmich.edu/reading_horizons)



Part of the [Education Commons](#)

## Recommended Citation

Henk, W. A., & Rickelman, R. J. (1992). Trends in Statewide Reading Assessment: A Closer Look. *Reading Horizons: A Journal of Literacy and Language Arts*, 33 (1). Retrieved from [https://scholarworks.wmich.edu/reading\\_horizons/vol33/iss1/6](https://scholarworks.wmich.edu/reading_horizons/vol33/iss1/6)

This Article is brought to you for free and open access by the Special Education and Literacy Studies at ScholarWorks at WMU. It has been accepted for inclusion in Reading Horizons: A Journal of Literacy and Language Arts by an authorized editor of ScholarWorks at WMU. For more information, please contact [maira.bundza@wmich.edu](mailto:maira.bundza@wmich.edu).



# **Trends in Statewide Reading Assessment: A Closer Look**

**William A. Henk  
Robert J. Rickelman**

In recent years significant changes in our thinking about reading have ushered in a much needed reform movement in reading assessment practices. The formal origins of this shift can be traced to exciting developments in statewide reading assessments that were rightly touted in the professional literature during the latter 1980s (Carbo, 1988; Illinois State Board of Education, 1988; Roeber and Dutcher, 1989; Wixson, Peters, Weber, and Roeber, 1987; Valencia, Pearson, Peters, and Wixson, 1989). Reports about truly innovative testing practices in pioneer states such as Michigan and Illinois seemed to signal the beginning of a national trend in reading assessment. In this paper, we report the results of a survey aimed at documenting the extent and nature of such a trend.

## **A new definition of reading and assessment**

The assessment reform movement gained momentum as educators began to question existing beliefs about the reading process (Valencia and Pearson, 1987). The idea that effective reading hinged on a large number of separate, specific, and measurable subskills gave way to a holistic definition in which facile reading involved the orchestration of a number of related strategic processes, all intended to help a reader create meaning from a text (Henk, in press; Squire, 1987). In this new scenario, reading represented a

dynamic interaction between the characteristics of a particular reader, the attributes of a specific text, and the unique context in which the reading occurs (ISBE, 1988; Wixson and Peters, 1984). This new definition made it clear that existing reading tests, by emphasizing mastery learning of small, discrete enabling skills, failed to reflect recent advances in reading instruction and research.

In Michigan and Illinois, statewide testing took a dramatic turn away from traditional models of reading assessment. Factors such as reader prior knowledge (topic familiarity), the effective use of before, during and after reading strategies, and school and home reading habits and attitudes played a role in large scale reading assessments for the first time. The new statewide testing formats also included full length, authentic narrative and expository texts drawn from children's magazines, tradebooks, literature anthologies, and content area texts. These passages represented quite a departure from the short, contrived passages used in existing statewide assessment instruments. Students were even asked about their interest in the passage and how easy or difficult they felt the passage was to read (Roeber and Dutcher, 1989).

Comprehension questions on these unique tests centered on higher level thinking processes and were framed around key structural elements and ideas in the passages (ISBE, 1988; Roeber, Kirby, Dutcher, and Smith, 1989). Passages were mapped using story grammar formats (setting, characters, problem, key events, resolution, and theme) and graphic organizers to make sure that questions focused only on important ideas.

Interestingly, at roughly the same time as these statewide assessments were taking shape, the National

Assessment of Educational Progress (NAEP, 1987) announced plans to measure reading performance in a similar manner. The NAEP tests planned to use a larger variety of reading materials, to assess complex processing of information, and to measure students' reading strategies and attitudes toward reading (Carbo, 1988).

### **Why take a closer look?**

There are a number of important reasons why educators need to examine trends in statewide reading assessments. As Valencia and Pearson (1987) and Afflerbach (1990b) suggest, reading achievement tests serve not only to assess student achievement and teacher effectiveness, but also to evaluate programs, and to group and place students. None of these functions should be taken lightly.

Moreover, tests have been instrumental in shaping current reading instruction (Anderson, Hiebert, Scott, and Wilkinson, 1985). Teachers tend to emphasize the skills, processes, and content that they expect to be on the test. Many believe that test publishers can better define what is important to teach, and so, they look to tests to inform curricular decisions (Valencia, Pearson, Peters, and Wixson, 1989). Consequently, if statewide assessments remain mired at the skills level, reading instruction will continue to be driven by outdated mastery learning models. Instructional progress could be severely limited if this incorrect model is used to define and assess reading proficiency at the statewide level. Likewise, inappropriate or insensitive tests can lead to abandonment of effective instruction (Valencia and Pearson, 1987).

On the other hand, if tests emphasize such productive strategies as surveying the test, making predictions, determining genre and text structure, setting goals, monitoring

and fixing-up comprehension, and summarizing and evaluating, reading instruction will tend to follow suit (Henk and Moore, in press). For this reason, assessments at all levels should attempt to tap elements like prior knowledge activation, text structure awareness, metacognition, and critical thinking. Unfortunately, states that assess reading in such an enlightened manner seem to be in the minority (Afflerbach, 1987).

Our goal was to determine if statewide reading assessment practices had changed appreciably in light of current theoretical and applied developments in the field. The study reported here builds upon the work of other authors who have addressed the general topic of statewide assessment (Bowers, 1991; Glassapp and Poggio, 1991; Shepard, 1989). In particular, it expands upon other recent, related surveys of statewide reading assessment practices whose aims were somewhat different (Afflerbach, 1990a; Dillingofski, 1990; Steel and Meredith, 1991). We were specifically interested in knowing if (and how) states assessed unique elements such as prior knowledge, reading strategies, and habits and attitudes; what kinds of reading passages and comprehension questions were used; how test data would be employed; and what plans existed for future assessments.

## **The survey**

To determine the status of statewide reading assessment practices, we developed a survey instrument that would be sent to state department of education officials in all 50 states. These individuals were listed in the Staff Roster of the Association of State Assessment Programs (Roeber, 1989) as coordinators, directors, or supervisors of statewide testing programs.

A cover letter explaining the study accompanied the survey instrument. Respondents were asked to reply to the survey within three weeks and to send supporting documentation if possible. When no response was received by that date, a second mailing of the survey and a revised letter was sent. If no response arrived within two more weeks, follow-up telephone inquiries were made. In some cases it was reported that the surveys had not been received, that the addressee no longer served in a coordinating capacity, or that another individual within the department should rightfully respond to the survey. We received verbal commitments from all states that completed surveys would be forthcoming. Ultimately, forty-nine states replied.

The survey instrument consisted of nine parts: 1) general statewide testing information; 2) information about reading passages; 3) test administration procedures; 4) comprehension assessment; 5) prior knowledge; 6) reading strategies; 7) reading habits and attitudes; 8) use of test data; and 9) future plans for statewide assessment.

In the first section we asked general questions about whether statewide testing in reading was done at all, and if so, what grade levels were assessed and what tests were used. Here we were trying to get a sense of what percentage of states assess reading at the statewide level, what grade levels tend to be tested, whether the test format remained the same across grade levels, whether certain standardized tests were used more often, and whether states took responsibility for their own test development.

The next section on reading passages asked about the use of narrative and expository texts and whether the passages were derived from real sources or were specially prepared for the test. Questions were also asked about

other text selection factors including passage length, interest, story grammar considerations, organization, and content area affiliation.

For test administration procedures, we wanted to know whether sample or practice exercises were part of the test format. Other items in this section inquired about: time allotments, the use of brief introductory material for schema activation, and the order in which major assessment elements (prior knowledge, passage reading, comprehension, reading strategies, and reading habits and attitudes) occurred during testing.

Section four on comprehension assessment queried respondents about the number and type of questions that were asked at each grade level assessed. Besides identifying the rough percentage of questions that fell into literal, inferential, and higher levels, we asked about the use of story grammar as a basis for framing questions about narrative passages; embedded versus post-reading questions, and other factors considered in questioning such as importance of ideas; the desired match between question and specific objectives; and relative ease of framing the question.

The prior knowledge portion of the survey first asked if topic familiarity was assessed at all. If it was assessed, then respondents were asked about the format (multiple-choice, open-ended short answer, extended free writing, vocabulary knowledge or other) and how the data were used.

The following section, which dealt with reading strategies, again checked to see if this type of assessment played any role whatsoever in the assessment, and if so, whether general strategies served as the focus or whether strategies

specifically tied to the passage tended to be highlighted. Next, respondents were requested to check off any strategies that were assessed from a thorough list of before, during and after strategies (e.g., surveying, predicting content, predicting text structure or genre, purpose setting, imagery, self-questioning and paraphrasing, other comprehension monitoring strategies, fix-up routines, summarizing, evaluating and studying).

The section on habits and attitudes was designed to indicate whether states asked their students about a range of school-based and home-based considerations. This part of the survey included several items that dealt with reading habits in school and at home, general attitude toward reading, students' perceptions of themselves as readers, purposes for which students might read, types of materials that are read, habits of others in the home, instructional techniques teachers use, writing-related activities, and students' willingness to discuss books and make use of the library.

Section eight asked how test data were actually used. Items centered on whether information from the tests were used for student placement, diagnosis, evaluating teaching effectiveness, certifying graduation eligibility, determining district funding, or comparing districts and schools.

The final section of the survey dealt with states' future plans for assessment. We were particularly interested in determining if states not currently assessing reading in a manner consistent with contemporary thinking about the reading process had plans to move in this direction.

## **The findings**

The information received from each statewide coordinator was used to categorize testing practices into the



sections described above. While data from 49 states was collected, there are not always 49 responses in each section. The reason for this is that some states allow several methods of assessing students, and the description of the test itself varies according to the choices of individual school administrators. In other states, the methods do not easily break down into one clear category. At times, none of the categories from the survey adequately described the test situation. At other times, several descriptors were needed. Therefore, even with simple yes/no type questions, the number of responses may either add up to greater or less than forty-nine.

The survey generated a considerable amount of data. For practical purposes, however, only the highlights are reported here. Since our aim was to discern trends, information about individual states is not presented. Specific information about other characteristics of individual states' testing programs that fell outside the scope of our survey can be found in Afflerbach (1990a) or Dillingofski (1990). Our major findings were as follows:

- Reading assessment currently occurs in 43 of the 49 states that responded. The most common grade levels at which assessment occurs are eighth grade (30 states), third grade (28 states), and sixth and eleventh grades (23 states each). In large measure, these results concur with Afflerbach's (1990a) findings. However, according to our survey, six more states test at the eleventh grade level. All but two of the states mandate the same test for all children at a particular grade level. These states offer individual districts several options for reporting test data.

- As expected, many states use national level standardized test subscores to assess reading. For instance, six

states use the reading subtests of the most recent editions of the following achievement tests: the *California Test of Basic Skills*, the *Iowa Test of Basic Skills*, and the *Stanford Achievement Test*. Twenty-two states use their own tests developed within the state. Some use these self-made tests exclusively, while others use them in conjunction with national level standardized tests.

- Information about reading passages indicated that 36 states test reading ability using both narrative and expository types. In nine states, the pattern varies by grade level, usually by including more expository passages for older students. Roughly half of the states include passages from real sources, while the other half uses passages specially written for the test instrument. About 30 percent of the states include graphic aids with the passages.

- Just over half of the states use short passages to assess comprehension rather than longer passages. Twenty-nine states chose passages because they were interesting, while 22 states chose passages because they conformed to a story grammar. Only 14 states purposely included content-specific passages, and these were evenly divided between science and social studies. Several states listed other sources for choosing passage material such as consumer labels, timetables, and newspapers. Overall, the two criteria that seemed to be the most influential in choosing passages were interest and length.

- All but seven of the states provided practice examples on the tests. Twenty-nine of the 43 states give timed tests, with the time period ranging from 10 to 50 minutes. Only 33 percent of the states provide an introduction to the content of the reading passages. Other than the fact that comprehension questions consistently followed passage reading,

there were no consistent patterns as to when prior knowledge, reading strategies, and habits and attitudes were assessed in the few states that did so.

- Two factors that varied widely from state to state were the number of passages students read to measure comprehension and the number of questions asked relating to these passages. The median number of passages per grade level ranged from 5.0 for twelfth graders to 10.0 for eighth graders, with an actual range from only one passage per assessment to 40 passages. The number of questions asked showed similar variance. The median number of questions asked for a reading assessment ranged from 22 questions for first graders to 97 questions for twelfth graders, with the actual number of questions asked each student ranging from five to 400 questions.

- The types of questions asked were almost evenly divided among literal (38 percent) and inferential (36 percent) types. Higher level questions occurred 26 percent of the time. All but four states placed the questions at the end of the passages. These states included questions embedded within the reading passages. The major factors considered in developing the questions (regardless of placement) were that they tapped main ideas and that they measured statewide objectives.

- Prior knowledge of passage content is assessed by only six states. Prior knowledge items tended to be multiple choice although vocabulary knowledge items were also reported. Responses about how the prior knowledge data were used varied considerably. Some states tied the score directly or indirectly to an individual's comprehension performance, while others were concerned with programmatic, school and district comparisons.

- Reading strategies played a role in the statewide assessments of 11 states. In 10 of these states, students are asked to predict content and to reflect upon or to evaluate a passage after having read it. Nine states ask students to identify text structure and genre and eight states use items that measure readers' surveying of the text prior to reading. Only half of the states attempted to tap metacognitive strategies such as the comprehension monitoring techniques of self-questioning and paraphrasing, and various fix-up routines such as rereading, reading ahead, using context, and asking for help. Strategy assessment is tied directly to specific passages in six states and is measured in a general fashion in only one state. The remaining four states use some combination of general and specific strategy assessment.

- Nineteen states inquire about the reading habits and attitudes of students as part of the assessment. These states seemed to be most interested in school and home-based reading habits, general attitudes toward reading, readers' self-perception, and types of materials read. They were less interested in teacher practices and the habits of other individuals within the home.

- The ways in which test results are used vary widely among the states. Twenty-seven states report using the tests for diagnostic purposes. Other uses include: district comparisons (20 states), student placement (14 states), funding determination (10 states), and evaluating teacher effectiveness (three states).

- Five states have no plans for assessing reading in the future. The remainder have plans to either implement or continue statewide reading assessment. Of the states that

will be testing, roughly one third expressed an interest in formats aligned with contemporary research findings and theories of reading. Specifically, plans were reported to develop items for prior knowledge, reading strategy use, and reading habits and attitudes and to include authentic texts when possible. The particular provisions for these assessments had not, in many cases, been determined.

## **Discussion**

According to Dillingofski (1990), there seems to be a positive trend emerging in statewide reading assessment practices. She reports that while 68 percent of the states use traditional standardized testing exclusively, fully 50 percent of the states already have some form of performance-based testing in place or have plans to change to this type of assessment in the near future. These tests are reportedly more holistic than their predecessors. The move away from standardized tests is indeed welcome, but the nature of the various performance-based tests that will be used is of critical importance. Unless these tests are genuinely more in line with current theories and practices of teaching reading, they may not represent a significant improvement.

On the basis of our survey, we are inclined to agree that a positive trend is occurring, but only in part. It is encouraging to note that one third of the states wish to remedy the mismatch between innovative instruction and traditional reading assessments. However, it is equally disappointing that two thirds of the states have no plans to move toward newer testing formats.

In our estimation, the trend toward enlightened statewide reading assessment is not nearly definitive enough. The typical state either uses an existing national standardized test or has created its own local standardized

version. Either way, the tests tend to be very traditional. Even when a local version has been constructed, it tends to resemble its national test counterpart. Shorter, inauthentic passages are used; no assessments are made of prior knowledge, reading strategy use, or habits and attitudes; and considerations of text structure and higher level reasoning are secondary at best.

In fact, only about 10 percent of the states presently qualify as being on the cutting edge of reading evaluation. These few states include provisions for the assessment of prior knowledge, reading strategy use, *and* reading habits and attitudes. Tests in these states are further marked by the use of authentic narrative and expository texts that are well structured and lend themselves to higher level thinking. Unfortunately, because so few tests tap prior knowledge and reading strategies, only a handful of states receive a fairly complete picture of children's reading performance.

It may be too soon to expect pervasive changes to occur in statewide reading tests. Despite the considerable notoriety of the Michigan and Illinois initiatives, the attention has been relatively recent. At the same time, though, dramatic changes are occurring in reading and language arts instruction nationwide (Monson and Pahl, 1991; Ridley, 1990) and these changes must be complemented by significant shifts in the way instructional outcomes in reading are measured (Henk, in press; Valencia and Pearson, 1987).

Our results suggest that at the statewide level, reading assessment has not kept pace with instruction. Apparently, the tests used to measure students' reading achievement are still primarily skill-based rather than broad-gauged assessments of reading ability. When instruction is based upon one view of the reading process and assessment is

based upon a contradictory view, educators and the children they serve both suffer (Anderson, Hiebert, Scott, and Wilkinson, 1985).

One related concern is that children will receive mixed messages about reading. Unless local, state and national assessment are in concert with one another, confusion will result. It would seem beneficial for state and national leaders in assessment to come together with the intent of establishing some common ground. In light of the current trend toward increased testing and accountability in all areas of education, such a consensus seems to be essential. If current trends toward testing are not consistent with theories driving the curriculum, and if these test scores are taken as measures of student achievement, a false perception of the reading ability and related problems in American education could easily result (Carbo, 1988).

Over the next few years, it will be interesting to monitor the changes that take place in statewide and national reading assessments. Hopefully, more states will begin to embrace the value of newer testing practices like those used in Michigan and Illinois. More states may come into the fold as better ways are found to measure prior knowledge and strategy use. Others may jump on the bandwagon as formats other than multiple-choice items are refined or when writing plays a greater role in responding.

Since this survey was conducted, there has been talk of creatively combining reading and writing assessments, getting children to integrate information across more than one passage, collecting individual oral reading samples, and even large scale portfolio assessment. All of these innovations are exciting, but the challenge of developing, administering and scoring these tests and convincing those

in power of their value rests squarely with literacy educators. We need to be up to the task.

## References

- Afflerbach, P. (April 1987). *The statewide assessment of reading*. Paper presented at the annual meeting of the American Educational Research Association, Washington DC.
- Afflerbach, P. (1990a). Statewide reading assessment: A survey of the states. In P. Afflerbach (Ed.), *Issues in statewide reading assessment*, 101-161. Washington DC: ERIC Clearinghouse on Tests, Measurement and Evaluation.
- Afflerbach, P. (1990b). The call for assessment of reading at the statewide level. In P. Afflerbach (Ed.), *Issues in statewide reading assessment*, 7-14. Washington DC: ERIC Clearinghouse on Tests, Measurements and Evaluation.
- Anderson, R.C., Hiebert, E.H., Scott, J.A., & Wilkinson, I.A.G. (1985). *Becoming a nation of readers: The report of the commission on reading*. Washington DC: National Institute of Education.
- Bowers, J.J. (1991). Evaluating testing programs at the state and local levels. *Theory into practice*, 30, 52-60.
- Carbo, M. (1988). What reading tests should measure to increase literacy in the U.S. *Phi Delta Kappa Research Bulletin*, 7, 8-9, 11.
- Dillingofski, M.S. (1990). States move toward new testing model. *AAP Reading Initiative News*, 1-2.
- Glassapp, D.R., & Poggio, J.P. (1991). Mandated competency testing programs: Evaluating state and local initiatives. *Theory Into Practice*, 30, 61-68.
- Henk, W.A. (in press). New directions in reading assessment. *Reading and Writing Quarterly*.
- Henk, W.A., & Moore, J.C. (in press). Facilitating change in school literacy curricula: From statewide initiatives to local implementation. *Journal of Reading*.
- Illinois State Board of Education. (1988). *Assessing reading in Illinois*. Springfield IL: Illinois Department of Education.
- Monson, R.J., & Pahl, M.M. (1991). Charting a new course with whole language. *Educational Leadership*, 48, 51-53.
- National Assessment of Educational Progress. (1987). *Reading objectives: 1986 and 1988 assessments*. Princeton NJ: National Assessment of Educational Progress.
- Ridley, L. (1990). Enacting change in elementary school programs: Implementing a whole language perspective. *Reading Teacher*, 43, 640-646.
- Roeber, E.D. (1989). *Staff roster of the Association of State Assessment Programs*. Lansing MI: Michigan Department of Education.
- Roeber, E.D., & Dutcher, P.A. (1989). Michigan's innovative assessment of reading. *Educational Leadership*, 46, 64-69.



- Roeber, E.D., Kirby, C.S., Dutcher, P.A., & Smith, R.L.C. (1989). *Michigan educational assessment program reading test blueprint*, 3. Lansing MI: Michigan Department of Education.
- Shepard, L. (1989). Why we need better assessments. *Educational Leadership*, 46, 4-9.
- Squire, J.R. (1987). A special issue on the state of assessment in reading. *The Reading Teacher*, 40, 724-725.
- Steel, J.L., & Meredith, K. (1991). Standardized measures of reading achievement or placement of students in Chapter 1 and learning disability programs: A nationwide survey of assessment practices. *Reading Research and Instruction*, 30, 17-31.
- Valencia, S.W., & Pearson, P.D. (1987). Reading assessment: Time for a change. *The Reading Teacher*, 40, 726-732.
- Valencia, S.W., Pearson, P.D., Peters, C.W., & Wixson, K. (1989). Theory and practice in statewide reading assessment: Closing the gap. *Educational Leadership*, 46, 57-63.
- Wixson, K., & Peters, C. (1984). Reading redefined: A Michigan Reading Association position paper. *The Michigan Reading Journal*, 5-6.
- Wixson, K., Peters, C., Weber, E., & Roeber, E. (1987). New directions in statewide reading assessment. *The Reading Teacher*, 40, 749-754.

*William A. Henk is a faculty member in the Department of Education and Reading at Penn State at Harrisburg, Middletown Pennsylvania. Robert J. Rickelman is a faculty member in the Department of Teaching Specialties at the University of North Carolina at Charlotte, Charlotte North Carolina.*

**Copies of READING HORIZONS themed Issues...**

are available for ordering. Volume 31, No. 5 (June 1991) is devoted to the subject of Reading Recovery. Guest editor for this issue was Dr. Jim Burns; Marie Clay provided the lead article. Articles in Volume 32, No. 5 (June 1992) are centered around the theme of Alternative Methods of Grouping for Literacy Instruction. Guest editor for this issue was Dr. Mary E. Hauser; the lead article is by Richard L. Allington.

Back copies may be ordered with the following prices:

1-5 copies	\$5.00 per copy
6-15 copies	\$4.00 per copy
16 or more copies	\$3.75 per copy

Add 10% for postage and handling; send orders to Business Manager, *Reading Horizons*, Western Michigan University, Kalamazoo MI 49008; include payment or institutional purchase order.